Private-by-Design: Towards Personal Local Clouds

Roberto G. Cascella, Christine Morin, Jean-Pierre Banâtre, Thierry Priol

Private-by-Design: Towards Personal Local Clouds

Roberto G. Cascella, Christine Morin , Jean-Pierre Banâtre, Thierry Priol

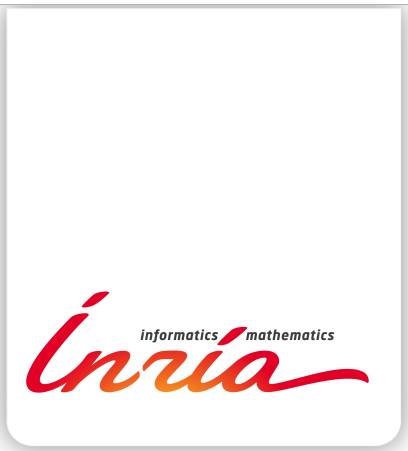
RESEARCH REPORT

N° 8634

November 2014

Project-Teams Myriads

ISSN 0249-6399 ISRN INRIA/RR--8634--FR+ENG



Private-by-Design: Towards Personal Local Clouds

Roberto G. Cascella*, Christine Morin *, Jean-Pierre Banâtre†, Thierry Priol*

Project-Teams Myriads

Research Report n° 8634 — November 2014 — 21 pages

Abstract: Recent trends of cloud computing for managing, storing, and processing data have made possible blending information from heterogeneous sources that is available at anytime and everywhere. However, recent large-scale information leakage of users' data have raised concerns in guaranteeing their privacy. Many users are not fully aware of the privacy risk in sharing their data, what information others can infer from the different data items, how data is handled by cloud providers and web applications, and how persistent sensitive data is.
In this paper we discuss the need to design new *personal* clouds with privacy guaranteed by design and we stress the importance of *locality* in scope and availability to guarantee privacy. We argue that users must always have control over their data without the need to delegate the management of their sensitive information to public cloud providers. We define the models that drive the design of a new system. Then we present our system along with the technical challenges and directions to reconcile security, usability, and privacy with the existing infrastructures.

Key-words: Cloud computing, private-by-design, security, privacy, personal cloud system, personal data management, home gateway, ephemeral mobile clouds

Email: {roberto.cascella, christine.morin, jean-pierre.banatre, thierry.priol}@inria.fr

* Inria, Rennes (France)
† University of Rennes 1, Rennes (France)

RESEARCH CENTRE
RENNES – BRETAGNE ATLANTIQUE

Campus universitaire de Beaulieu
35042 Rennes Cedex

Privé par conception : vers des clouds personnels locaux

Résumé : Des évolutions récentes dans le domaine des nuages informatiques pour la gestion, le stockage et le traitement de données ont rendu possible le mélange d'informations provenant de sources hétérogènes qui sont disponibles partout et à tout moment. Cependant, les récentes fuites massives d'informations concernant les utilisateurs de nuages soulèvent des inquiétudes liées à la protection de la vie privée. De nombreux utilisateurs ne sont pas complètement conscients des risques inhérents au partage de leurs données, de l'inférence d'information à partir du croisement de données, de comment les données sont gérées par les fournisseurs de nuages informatiques et d'applications web, et de la durée de conservation de données sensibles.

Dans cet article, nous discutons le besoin de concevoir de nouveaux nuages informatiques personnels pour lesquels la protection de la vie privée est garantie par conception et nous soulignons l'importance de la localité en termes de portée et de disponibilité pour garantir le respect de la vie privée. Nous affirmons que les utilisateurs doivent toujours garder le contrôle sur leurs données sans avoir besoin de déléguer la gestion de leurs données sensibles à des fournisseurs de nuages informatiques publics. We définissons les modèles qui guident la conception d'un nouveau système. Nous présentons ensuite ce système ainsi que les défis techniques et les directions de travail associés pour concilier sécurité, facilité d'utilisation et protection de la vie privée dans le contexte d'infrastructures existantes.

Mots-clés : Nuages informatiques, privé par conception, sécurité, protection de la vie privée, système de cloud personnel, gestion de données personnelles, passerelle domestique, nuages informatiques mobiles éphémères

1 Introduction

A multitude of smartphones applications and services are flourishing due to the need of nomadic users to be guided in their day to day life while being always connected. The day life is characterised by the presence of sensors everywhere generating a continuous flow of data that is processed to generate information used to better serve the users, to ameliorate their quality of life, and to tackle societal challenges. Nowadays, citizens are not only data and service consumers, but they produce data and take active participation in tasks that require their manual intervention. We live in a world that could be seen as an immense place where data generated from different and heterogeneous sources are blended and are potentially accessible at anytime and everywhere. This is facilitated by the reduced cost in storage and computation available everywhere.

However, this vast amount of data might contain and reveal personal user information, entrusted to service providers who can now identify and create an accurate profile of the users. Moreover, personal information of a user can expose others without any means for them to prevent this [1].

In this world of big data, users are generally concerned of the risks in sharing personal data and are reluctant in trusting organisations and service providers. The recent large scale data leakage [2] has contributed to a socially driven change in the factors that influence the users entrust of personal data. More than 78% of users find it hard to trust how companies use their personal data, perceiving that they hold too much information [3]. The shortcoming of trustworthy tools to verify how organisations handle data and the increasing value users give to personal data contribute to this mistrust in organisations.

In the "big brother" era, cloud computing plays a key role as it is more and more used by companies to provide their services, including mobile and Internet of Things (IoT) applications that use the cloud as backend. The low cost in storage will incentivize companies in collecting and retaining permanently users' data for potential further use. The offloading of services and computation to third parties to cut down the service costs can also make hard for the user to track and retain the control over her data.

This paper advocates the need to design a new system for personal clouds to regain control over personal data. The *Personal Cloud* is the abstract view of the user in the cloud computing domain and we refer to this term to indicate user-centric cloud technology, while in the literature it has been used with a different meaning to identify storage services for personal data. We enforce a private-by-design proactive approach [4] to anticipate privacy risks and guarantee that appropriate control over user privacy is always maintained during the lifetime of the system. In our private-by-design system, privacy should also come at no cost over the usability of the system, by not creating any disruption or limiting the actions of the user, who should be able to state policies over her data. Then, the system should guarantee that privacy requirements are automatically enforced without requiring specific actions by the user. The key is to push data *locally* to the user, under her direct control, where access policies can be enforced and dynamically adapted based on the sensitivity of the information or on the situation. A context-aware data usage also becomes of outmost importance as data value changes based on the situation [5].

Our contribution is the definition of the user, application, and adversarial models that drive the design of a new system and a qualitative analysis of the existing architectures under these models. Driven by the outcome of this analysis, our last contribution is the proposition of a new architectural system design and a discussion of the technical challenges to pave the way toward such a privacy aware system.

This paper is structured as follows. Sec. 2 deals with the assumptions and models driving the system design. Sec. 3 summarises the existing trend toward a user-centric approach and Sec. 4

presents the private-by-design system. Sec. 5 lists the open challenges for implementing such a system and analyses the available technology. Finally, Sec. 6 concludes the paper and discusses future directions.

2 Models and assumptions

Our system main objective is to bring trustworthiness over undependable cloud providers by introducing new practises for users to keep a direct control over their data, like pushing the content local to the user. In this section we first give our definitions and then discuss the models that are used throughout this paper to analyse existing solutions and validate our system design.

Privacy refers to the right of retaining control over personal data, what information is shared with whom, and how it is used to prevent its abuse. Herein, we adopt the very broad definition of *personal data* reported in the European Directive 95/46/EC which groups any information relating to an identified or identifiable person [6]. We further distinguish between personal identifiers and data. Personal identifiers refers to user identity or any other credential used to access a service that allows to identify the user. On the other end, data, when processed, generates information that allows to profile the user (e.g. her habits, movements, or opinions) and to infer additional knowledge that the user would like to keep personal (e.g. network of contacts, location, or work related data). In the rest of the paper, we use the term personal data to indicate this last type of information and simply identifiers or identity for the former.

Another important concept related to privacy is the *sensitivity* of the information, useful to assess data privacy levels, as not all the information requires the same level of protection. The sensitivity is subjective as data value changes based on the situation, and could be assessed using crowdsourcing solutions as proposed in [7]. Factors that contribute to determine sensitivity levels are the user preferences and the environment in which the data is shared, e.g. the value of the service in exchange or the trust in the service provider [5]. This can be summarised in defining the *context*, which could be metadata associated with the data, useful to assess the level of privacy.

User-centric personal data ecosystems are a viable way to manage privacy levels where individuals can have the full control over their personal data and how they are used. The trust in service providers needs to be well substantiated with actual practise of data protection that can be easily verified and evaluated by the users in exchange of the access to valuable services. Context-aware data usage and fine-grained policies under the control of the users are two independent ways to achieve the sustainability of this ecosystem. The usage of data should be limited to the context where it was initially collected or later revised by the user. However, new policies should be enforced to allow the user to manage personal information and decide who has the right to access which information and for what purposes. While they are different in the way they are conceived, they are fundamental and can coexist to address data privacy and reduce the risk of information misuse.

In this paper we focus on the design of a system that will facilitate the emergence of this user-centric personal data ecosystem that leverages the existing cloud infrastructure to guarantee users privacy regardless of data sensitivity levels. In our model, we do not set an *a priori* definition of sensitive information due to the heterogeneity of the shared data items we consider. Instead, we repute all information valuable: the sensitivity is just a matter of the specific privacy policy associated with the content.

Table 1: User and entities trust relationship

Entities / **User**	*Service providers*	*Selected entities*	
	Cloud	**Acquaintances**	**Contract-based**
No trust	–	–	–
Content based trust	Limited to service access and usage	Limited to some content	Limited to the context
Selective provider trust	Crowd anonymity and provider reputation	Content based and peer reputation	Context dependent
Don't care	As long as they access the service	Trust by default	

2.1 User model

The private-by-design approach we advocate in this paper puts a lot of emphasis on the role of the users. The needs of the individuals and how they perceive the risk for their privacy are relevant to understand which challenges and obstacles the technology should overcome for a wider adoption. We group the users in four different behavioural categories based on their willingness to share content and the trust they pose in others, i.e., users and public cloud providers, in exchange of valuable services.

Table 1 summarises the trust relationships among the users and entities, divided in two main categories: cloud service providers, such as Infrastructure as a Service (IaaS) and Service/Storage as a Service (SaaS), and selected entities, indicating whether there is a direct relationship (acquaintances) or they collaborate in a given contract. The scope of the trust is limited to managing user data, e.g. file storage, data usage or handling personal identifiers to access a service.
The users are categorised as follows:

2.1.1 No trust

This user does not share any type of content with others. All data is kept locally and she does not use any public service for personal storage, nor she shares personal data or participate in collaborative frameworks or social networks. The result is like being wrapped in a *digital cotton wool* with no contact with the digital external world.

2.1.2 Content based trust

This user trusts the service providers or external entities with data only contextually relevant to the service. In this category we group users giving access to their position for geolocation services, or using storage public clouds for keeping content backup or sharing files with their peers.

2.1.3 Selective provider trust

This user trusts only a limited set of service providers for managing identifiers and personal data (e.g. content and emails). In most cases the trust is built on social perceptions and public acceptance in the community. Examples include well established public cloud providers, widely used social networks or email/web service providers. In other cases, trust is built on commercial

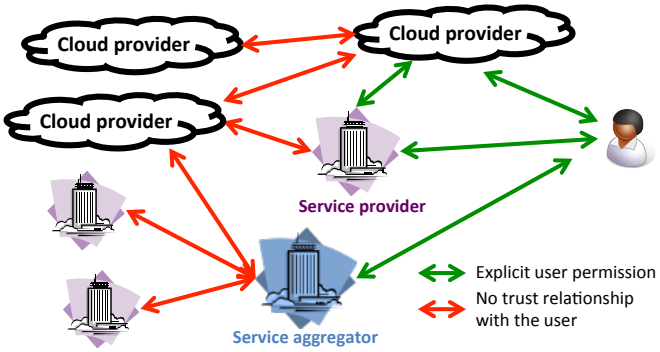


Figure 1: User-centric ecosystem: the green lines indicate user explicit permissions to handle personal information in exchange of valuable services; the red lines represent third party organisations collecting information anonymously or under the responsibility of the primary service provider.

relationship or the impact of the service provider in the public market. The acceptance of the risk of exposing personal information is based on the *wrong* assumption that the provider will not have any interest in the specific data of the user. Users lack knowledge and ability to control further use of that information by external entities with whom they have an indirect interaction. Fig. 1 depicts a user-centric ecosystem where users' personal data is shared at different levels with different entities, potentially having business relationships. Traditional thinking lays on the assumption that users can hide in a crowd that is too big to search through (*crowd anonymisation*) or that cannot be identified or tracked in the crowd [8].

2.1.4 Don't care

In this category we include users who don't care or are unaware of the privacy risks. They only want to access services and do not pose much attention on how their data is handled by service providers. Most of the users fall in this category. The reason is that security and usability are not always integrated in system design; for instance most of the time the privacy settings indicated by the service providers are either too complicated to be followed or a more fine-grained control over them limits or precludes the use of the service.

In this paper we address users in the last three categories. The goal is to build a system that preserves user privacy and increases the trust level by bringing data under direct control of the user, thus, building a trustworthy infrastructure on top of undependable public cloud providers.

2.2 Application model

Users rely more and more on mobile devices like smartphones or tablets for their daily activities, to manage their personal contacts, emails and photos, or to access services, e.g. geolocation based. These devices are equipped with hardware that can generate a continuous flow of data, like GPS or sensors and have also embedded computation capabilities to process data locally, being the end-points for user communications. Multiple application models can be defined based on the primary objective, whether to reduce the energy consumption, or the latency by performing the computation locally, or the privacy by keeping data in the proximity of the user. The reader can refer to [9] for a detailed survey and references to implementations.

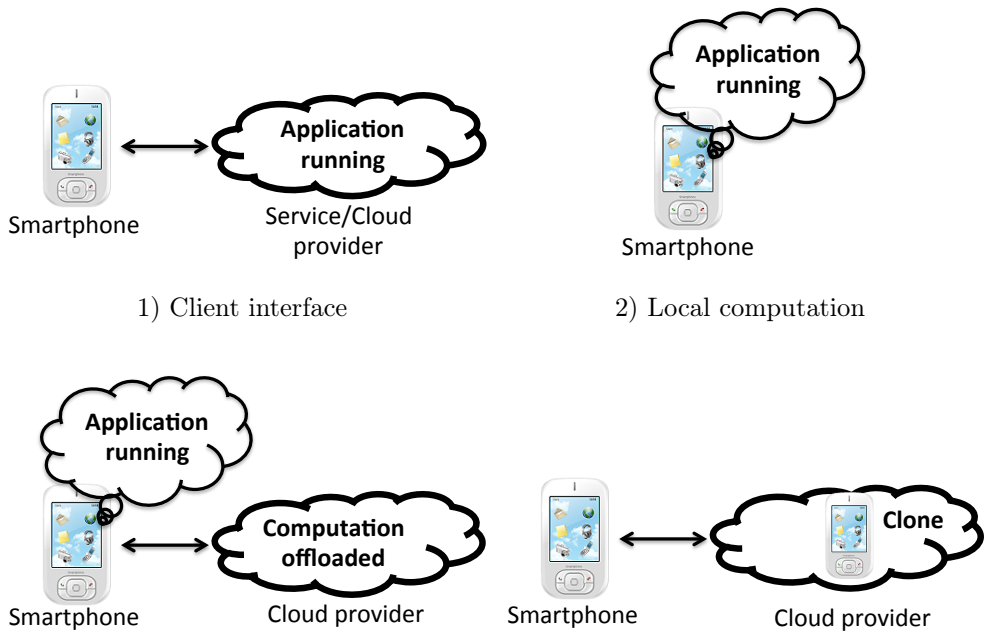


Figure 2: Application models.

For our purposes, we identify four different application models characterised by the way the infrastructure is used, as depicted in Fig. 2. Ad hoc mobile cloud is discussed as special case of no infrastructure usage for computation offloading.

2.2.1 Client interface

The mobile device is the interface with the user. Hardware collected raw data is transmitted to the service provider, that processes it and returns the result as part of the service to the user. Raw data is transmitted to the service provider where the computation is performed. Hence, the latter has full control of the user data, potentially exposing the user to privacy risks. Examples are geolocation services.

2.2.2 Local computation

The mobile application is running on the end-user device and the computation is performed locally. Examples are off-line applications, which do not interact with the infrastructure and do not require to synchronise data, stored in the device.

2.2.3 Computation offloading

A mobile application can decide to offload part of the computation to the cloud based on the network connectivity, the type of task or the type of data to be processed. The decision can be taken dynamically (elastic application partition), based on the required and available resources, or statically, already predetermined by the application programmer. The operation requires the

partition of the application into manageable (processed internally) and offloaded components (processed in the cloud).

Ad-hoc mobile cloud applications falls in the *computation offloading* category, where the cloud provider is other mobile peers and applications run on a group of mobile devices acting as a cloud. Each client provides access to computing resources, storage or Internet connectivity to other mobile nodes. This model becomes interesting for situations with low or no Internet connectivity to the infrastructure and service providers. Examples are collaborative and data sharing applications or those requiring user participation like in crowd computing applications. For instance, the user might decide to run the computation on the sensitive data locally, and then authorise the service provider to use part of the information or share it with other peers for ad-hoc mobile applications.

2.2.4 Clone cloud

A clone of the device is instantiated in the cloud. Applications and data could be partially or totally offloaded to the clone. The clone has the advantage of being potentially used for backup and recovery. However, this model requires fine-grained synchronisation techniques to keep the consistency of data and it is exposed to high privacy risks since an adversary could take control of the clone and access sensitive information without tampering the physical device.

The objective is to address all application types and in particular reduce privacy risks for those that share either raw data, offload computation or create a mobile clone.

2.3 Threat model

Clouds are big black-boxes, where the internals are not exposed: clients have no means to directly control how their sensitive data is handled. We consider a honest but curious adversary being interested in collecting and inferring more information about the user or potentially misusing the data for other purposes, e.g. advertising. Thus, we exclude potentially malicious users who tamper the system to collect or manipulate confidential data and focus on the way collected data is used without risks for privacy.

The adversary can be the service provider with whom the user shares voluntarily the information (explicit permission), third parties that can access users personal data anonymously, either to provide outsourced services or for other purposes, or cloud service aggregators, which integrate multiple cloud services (SaaS) into a single service, see Fig. 1. Herein, we identify a number of potential threats related to public cloud providers or services using cloud computing as backend, while more threats can be considered based on the context of the application. We exclude all potential threats that involve using information voluntarily and publicly shared by the user.

2.3.1 Identification threat

The access to public services is granted upon authentication of the user who discloses her identity. However, service providers might collect additional information that is not relevant to the service either during the authentication phase, e.g. the user location even if no geolocation services are needed, or from the content stored at the service provider. For instance, the adversary can use this information to infer additional data like working place of the user or habits. Some services today track user behaviour for a range of purposes, from sending targeted advertising to improving services.

If the cloud provider subcontracts services to third party clouds or integrates multiple services

(see Fig. 1), these external entities could potentially infer user personal information and determine the identity by correlating the shared information despite the user has not granted explicit permission.

2.3.2 Profiling threat

The adversary might receive information from different sources, i.e., the user or other service providers with whom both the adversary and the user have direct business relationship. Even if this data is anonymized, the adversary can infer the user profile by means of big data techniques [10]. Examples include profiling user behaviour or consumption habits in case of smart homes, user medical and financial situation, or business related sensitive information.

2.3.3 Storage threat

Cloud providers store and collect data on behalf of the users as part of the service. Usually they replicate the data across multiple systems (even tapes in some cases) and sites to increase the availability or to reduce the latency of accessing the data from multiple locations. When data should be automatically deleted at the end of the process or if the user asks to delete it at runtime, she has no means to verify that data is securely destroyed in a timely fashion. Moreover, if the cloud provider subcontracts part of the service to third party clouds, the user has no means to verify that data will still be securely retained.

We have limited our analysis to data management and not considered threats that could involve the operating system of the user device, the hypervisor of the physical machine hosting user instances, and the software used to manage and store all personal information. For the sake of system design we assume the software to be trustworthy, more is discussed in Sec. 6

3 Existing trends toward the network edge

The first solutions to provide cloud computing services refer to a centralised physical infrastructure to simplify the installation and maintenance expenses at the cost of being a single point of failure and increasing the energy consumption. The evolution of computing technology has been driven by many factors such as the need to reduce energy consumption or to accommodate the increased demand in processing power. Recent trends in industry have seen the creation of multiple data centers geographically scattered and interconnected with fast private networks to better serve the user needs and cope well in case of power outage, network failures, or natural disasters. In this section, we analyse the existing solutions and concepts related to cloud computing technology from the classical centralised approach, moving then to the concept of fog computing, which pushes the data computation in the network, and to mobile cloud which involves end user devices at the network edge. Fig. 3 depicts this technology trend.

Cloud computing technology comprises a large fraction of the services nowadays provided on the Internet both to final users (either directly or indirectly benefiting) and companies. Different abstraction models exist from the lower virtualised infrastructure (IaaS), to higher service solution stack (PaaS), ending with storage or software (SaaS) services. The centralisation of these services on large data centers relies on *public cloud* companies for the management of data and computing operations. *Private cloud* solutions might flourish thanks to new hardware technology at low cost and energy consumption, the availability of open source solutions to eliminate the costs of license, and more important because of the need to keep data locally and address security concerns: data being an important asset for the company.

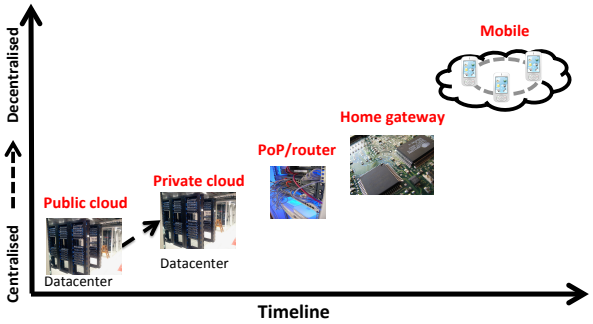


Figure 3: Technology trend for cloud computing: from the center to the edge of the network.

This trend has motivated the definition of new and extensible Internet architectures [11] and ecosystems to cope with the need for a reliable access to content and data, including cloud computing resources, with security being an essential feature. XIA [12] architecture design transforms the Internet from the classical host-to-host communication into a data and content centric paradigm by introducing the notions of principal, the originator or destination of a packet, including host, service, and content. This notion could be further extended to support any type of resource, application or usage model. The main focus is on security by using self-certifying identifiers for all the principals and including functions for trust management. Nebula [13] is a cloud computing -centric architecture addressing specific requirements on the network for dependability, security, flexibility and extensibility to create a utility like service, for both storage and computing. This requires new data control plane and networking techniques to interconnect data centers while guaranteeing dependability and security in both access and transit. Extensibility is another important property of the Nebula architecture to address future connection needs of new applications.

To cope with a more reliable access to data and with the increased demand of utility computing and services with high performance and low latency, the computation and storage has been pushed at the edge of the network. One of the noteworthy initiatives is the Nano Data Center distributed computing architecture reusing the ISP gateways to provide computing and storage capabilities [14]. This approach targets content and service delivery aiming for reduced energy consumption, where the *home gateway* is still under direct control of the ISP. In this same direction, the nanostores proposal collocates the computation with persistent data store [15], thus pushing data local to the user. Energy saving is also motivating a decentralised approach when certain applications can be off-loaded from centralised data centers for a given location and type of access network [16]. A recent initiative is Discovery which has locality as primary concern [17] by leveraging networking facilities, like Internet *Point of Presence (PoP)*, to deploy locality-based utility computing platforms. The aim is to exploit the bandwidth available in the backbone of the network to address the explosion in demand for bandwidth in the current generation of cloud computing, and to provide fast interconnection of storage and computing facilities, while offering a more reliable and fast access to cloud services.

Fog computing [18] is a new emerging concept as a result of the device ubiquity to access cloud services and the need to use more efficiently computational resources for a more scalable management of the network and services. Among the claimed benefits of fog computing is the push of the clouds at the edge of the network and the trend to keep the data in the network without relying on centralised services. Big data and the increased interest in Smart cities and

Internet of Thing (IoT) technology will facilitate the emergence of this new computing paradigm by keeping data and the computation geographically distributed and close to the end-users. Very interesting are also the new trends to build nano data centers for low computation applications using the ARM-based microprocessor [19], both at academic [20] and commercial [21] scale.

The increasing number of mobile applications has steered the interest toward *mobile computing*. The Mobile Backend as a Service (BaaS) model is emerging to link mobile applications with cloud backed storage & computing and cloud services in general. This approach, while serving mobile clients, relies on a centralised infrastructure and does not leverage the computational capabilities of mobile nodes. MobilityFirst [22] is a mobility-centric architecture proposal with the intent to address several challenges toward trustworthiness, location privacy, and usability with simple APIs that simplify application development. An interesting functionality is the use of in-network storage for packets in transit, as a form of generalised DTN-like routing, and of computing capabilities at routers, to make the architecture tolerant to nodes' mobility.

3.1 Models' based analysis

The approaches presented in this section delineate two important trends: (i) create a data-centric ecosystem that leverages the in-network capabilities and pushes the content and data toward the user and (ii) address security at all levels from communication to content. The future Internet architectures discussed above focus more on securing the system by ensuring that the parties' identifiers are certified or by certifying the communication path, or as for the MobilityFirst proposal on the protection of user location.

The application models discussed in Sec. 2.2 can be all implemented in the current Internet architecture at the cost of exposing the user to potential privacy risks if security mechanisms are not enforced. XIA, Nebula and MobilityFirst architecture proposals or the Discovery approach can definitely enable better performance by reducing the communication latency or facilitating the integration of security mechanisms. This is more pronounced for the case of CloneCloud or computation offloading where the network bandwidth might be more an issue.

If we consider the adversarial model discussed in Sec. 2.3, ad hoc solutions can be put in place to circumvent the potential privacy risks, but they are not integrated by design. Personal data might still be under the control of the service provider and a consistent enforcement of policies on data and communication might be very complex to achieve.

In our view, we introduce privacy as third dimension for the design of a user-centric ecosystem and focus more on a comprehensive approach, that could be implemented in the existing Internet architecture, and flexible enough to adapt to new ones. We believe in a more decentralised cloud model that extends the Fog computing paradigm to the extreme edge of the network, local and under the control of the user to preserve data privacy.

4 Privacy meets locality: our design approach

We push further the fog computing approach and propose to deploy personal servers to extent cloud infrastructures to the extreme edge of the network (e.g. at home). On the one hand this will leverage the in-network device computation capabilities. On the other hand this will increase the quality of experience of the users who will have better control of their data.

Our private-by-design cloud system makes privacy an essential component by embedding it into the design of personal cloud systems. The result is a user-centric and trustworthy ecosystem that leverages the existing but undependable cloud offer. Fig. 4 shows the high level architecture of our system and the interactions among the components. *Personal Data* means sensitive data

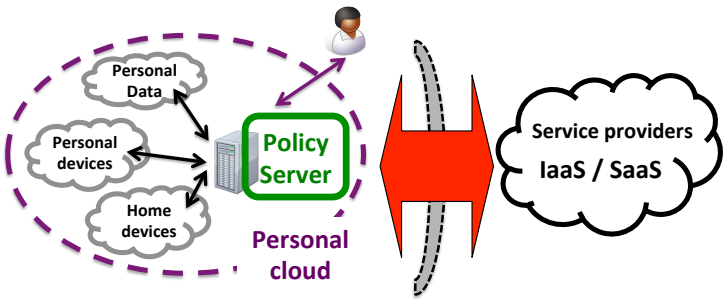


Figure 4: Private-by-design: Personal Cloud using existing cloud services (XaaS).

relating to the user, that if processed can either identify or profile the user; content like photos or data collected with home or personal devices fall in this category.

The *Personal Cloud* is the abstract view of the user in the cloud computing domain. In this paper, the term *Personal Cloud* refers to user-centric technology, while in the literature the term has been used with a different meaning to identify storage services for personal data. A Personal Cloud system does not replace cloud providers, like several open source projects or commercial products propose to do, but instead tries to complement the existing cloud offers by providing a better control on the treatment and storage of personal data. The Personal Cloud embodies a *policy server* that has dynamic fine-grained policies specified by the user to manage external communication, data sharing, and application partition components. Privacy is enforced at all levels. Data is kept local in the personal cloud and protected under direct control of the user who can decide with whom sharing data and which services to use by specifying policies for type of content or provider. Automated policies could be determined based on the sensitivity of the content [7] and the context [5]. For instance, the communication with certain services could be in clear, encrypted, or anonymised by using TOR-like mechanisms.

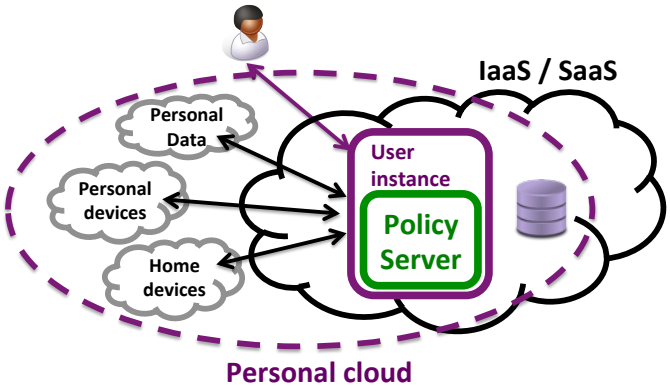


Figure 5: Private-by-design: Personal Cloud including instances and storage on cloud providers.

A *Personal Cloud* consists of the hardware and software components running on user equipment or as virtual instances run by the users on public clouds, for computation and storage, always under the control of the user, as shown in Fig. 5. In such a scenario, the user is able to leverage the existing undependable cloud infrastructure by running dedicated services, like the

policy server, to protect her privacy.

By pushing data and computation at the edge of the network, "locality" allows the Personal Cloud system to specifically address the needs of the users identified in Sec. 2.1, who can benefit public cloud services without putting their privacy at risk. The users continuously have control of their data, which is shared in a consistent way with external entities thanks to the role of the policy server.

As for the application models identified in Sec. 2.2, the Personal Cloud approach facilitates their deployment respecting the privacy of the user. In the Client interface model, the access to services and all communications go from the mobile device to the Personal Cloud which acts as intermediate point with the providers. The personal identifiers for the access and the data shared with third parties are under direct control of the *Policy Server* who might decide to either process the data locally, mask the information or share it without obfuscation. More interesting for our approach are the computation offloading and the Clone cloud models. In the former, the computation might be offloaded to the Personal Cloud or in the latter the mobile clone could be instantiated in the Personal Cloud either on the user hardware equipment or in the virtual instance the user runs on trustworthy public clouds (see Fig. 5).

Thanks to the policy server, which mediates the communication of the users with service providers, the identification and profiling threat could be thwarted and the persistent storage attack described in Sec. 2.3 could be limited to not sensitive data, offloaded to public clouds under user control.

In the remainder of this section we present three instantiations of this general architecture highlighting the implementation and the benefits of adopting the Personal Cloud approach for cloud based infrastructure and mobile services.

4.1 Home gateway cloud

One approach to implement Personal Clouds is to operate a physical machine at home under direct user control (Fig. 4). Such an approach has always been possible in theory using personal computers. However, energy consumption, and thus energy cost prevented it to practically happen. Rapid evolution of smartphone technologies is changing the situation very quickly. It is nowadays feasible to design very low power computing systems using smartphone components (such as ARM based processors), which are already available on the market (Raspberry, Cubieboard, ...). Such machines consume few watts (typically 3-5 watts), while providing very good performance when using multicore processors and thus can be always-on with an energy cost below 10$ per year.

A Personal Cloud, based on such a machine, can act as a home gateway between *personal devices* (e.g. smartphones, tablets, smart home appliances, ...) and cloud services wherever the user location is (inside or outside and far away of her home). All communications from these personal devices can be rerouted to the personal cloud before transmitting data to services hosted in untrustworthy cloud systems. The Personal Cloud hosts the policy server and can decide either direct communication without any filtering, anonymisation to avoid communicating the user IP address to the cloud service providers, or encryption to prevent cloud service providers to get access to personal data.

The policy server can be implemented using *man in the middle* proxy techniques to allow existing applications to take benefit of the proposed approach without requiring modifications. However, this deployment requires a good Internet connectivity at home to avoid adding too much latency. The deployment of FTTH or even FFTB is progressing or is planned within a matter of years in some countries. If the home Internet connectivity is not adequate, then the Personal Cloud can be instantiated in a virtual machine hosted by an IaaS cloud provider as

shown in Fig. 5.

4.2 Persistent cloud

The second approach is to build a Personal Cloud in a distributed way where trust relationships exist among the participants. Users are participants in a community or collaborative project and share their data and computational capabilities to solve complex problems. One way to implement such a system is to create *virtual organisations* controlled by an organisation or even self-organised, i.e., individuals voluntarily participate and manage the community. Trust relationships are built leveraging the organisation or direct physical interactions in self-organised systems. For instance, taxi drivers can create a community to share traffic information, client demands, and location data to better serve customers, or even offer infotainment services. Another example is friends sharing a flat or a family who can leverage the individual storage and processing capabilities to create a local cloud on top of their home gateway clouds.

A P2P Personal Clouds can be even created by taking advantage of the social networks that express in some ways relationships and trust between users. Social relationships have already been proposed in the past to enable friends sharing storage resources in a social network via incentive mechanisms [23] or computing resources in a social graph [24].

In both cases, the community of individuals participating in this decentralised system is *persistent* or does not change frequently. Such an approach exploits the individual computation and storage resources of the participants and relies little or barely on the infrastructure (see Fig. 6).

Persistent personal clouds can be implemented as a Grid of Personal clouds. Communities can be managed in a way inspired from the virtual organisation concept in science Grids [25] where users belong to organisations. However, in the context of persistent clouds resources are those of the participants with limited or no system administration skills. Thus, in persistent clouds not only privacy but also security mechanisms need to be implemented to control the access to shared devices and data. Reputation mechanisms can also be used to rank and trust the peers offering similar services.

4.3 Ephemeral clouds

User mobility coupled with the sensing and the wireless communication capabilities of their wearable and pocket mobile devices enable a wide range of novel application taking advantage of and for the benefit of user communities (or crowds). Examples of those mobile applications are mobile social networking applications enabling people to share content and information with their friends wherever they are and crowd-* applications for collectively solving problems affecting an individual (identify a free parking spot) or a community (resorbing a traffic jam, avoiding crowd public transportation vehicles, emergency situation following an accident or an earthquake) or for voluntarily participating in collective projects (participatory science).

The use of Personal Clouds enables the formation of *ephemeral clouds* where users can give access to personal data and computation capabilities to other peers without relying or little on the cloud infrastructure (see Fig. 6). Ephemeral clouds can now also spontaneously mushroom to serve the needs of the users for scope and time limited applications without the burden of having a dedicated infrastructure. This can lead to address the requirements of mobile applications leveraging the computing capabilities of the peers and the social structure to solve complex problems or simply to share information among a community of users. Ephemeral clouds will use the Personal Cloud system for supporting emerging mobile social and crowd-* applications in which privacy is a major concern.

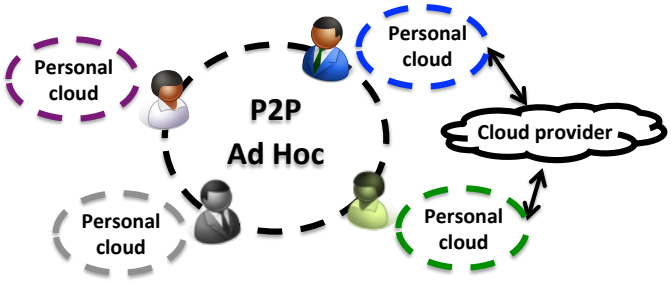


Figure 6: Private-by-design: P2P and ephemeral context-based cloud solution. Users form ad-hoc and P2P clouds to share data or to solve collaborative computational problems. Even if not required, public cloud services could be used to help in computation tasks or for temporary storage via the Personal cloud of some users.

Ephemeral clouds are characterised by the fact that they particularly (but not only) target mobile device users on the move who may opportunistically collaborate with peers exploiting the wireless communication and sensing capabilities of their mobile devices. They are highly dynamic as participating devices may join or leave at anytime depending on the user mobility and the wireless network availability. Moreover, ephemeral clouds are *ad-hoc clouds* by essence, thus they may be opportunistically created and exist for a limited period depending on the application or problem to be solved. Ephemeral clouds are different from persistent clouds because in this case no trust relationships exist between participants, or reputation schemes can be hardly put in place due to their high dynamic and ephemeral nature.

For instance, in participatory journalism, only people participating in a given event will be active and only during the event period. In such an application, it should be possible to check a contributing user location (for quality of the information) but user identity and location should not be kept and displayed along with the contributed data if the contributor wants to preserve her privacy. Another application is location of a missing or wanted person where all people present in a given geographical area should be able to contribute in an anonymous way if they wish and to control which peers and organisations get their data. Moreover, in applications such as the “person wanted" application, data should be deleted as soon as the problem is solved. Some applications may require data processing services to be off-loaded, for example due to the lack of resources (memory, computing, power). The ephemeral cloud system will be able to enforce the users’ resource sharing policies.

Ephemeral clouds are open and dynamic self-organised communities that last the time of the collaboration with peers (which can be few hours or days). People may opportunistically collaborate with people they did not know before the situation leading to the collaboration or participating mobile devices may be disconnected from the Internet. Thus, in ephemeral clouds not only privacy but also security mechanisms need to be implemented in a distributed fashion to control the access to shared devices and data at all time even during disconnection from the infrastructure network.

5 Technical challenges and directions

There are several technical challenges that academia and industry need to tackle when designing a system where privacy assurance is the default mode of operation. The objective will be for

the individuals to regain control over personal information and for service providers to gain a sustainable competitive advantage while being compliant with the regulations.

The considered underlying infrastructure is highly heterogeneous. Electronic devices are bought from diverse manufacturers that may have their own proprietary interface, and smartphones, computers and virtual machines run different legacy operating systems. This is not a new problem but it needs to be addressed to efficiently exploit device capabilities and the cloud service offer. Another interesting challenge is network connectivity. The Personal Cloud could be the bottleneck of our system being the gateway for all user interactions with external entities and services. On the other end, the use of inner computational capabilities and data locality will reduce the communication burden as data need not to be sent to centralised data centers. To ease the development of the system and application services running on top, a communication service is required to mask the complexity of the underlying networks due to the heterogeneity and unreliability, and to account for the remaining battery of mobile devices. In this context, P2P and Mobile Ad Hoc Networks (MANETs) will play a central role to support opportunistic collaborations.

In the remaining of this section we identify and focus on the challenges inherent to achieve private-by-design clouds.

5.1 Data management and accessibility

In the envisioned system architecture, data is produced in various streams, stored and accessed in the context of a swarm of personal devices interconnected with clouds providing a wide range of services. Different kinds of data are to be managed, with each type coming with various meta-data. A fine-grained management of meta-data is needed to be able to remove sensitive meta-data information, if this is a requirement stated by the user. Data management also needs to cope with this personal data tsunami by providing services and mechanisms for storing data, ensuring data accessibility anytime and anywhere, managing data time-to-live and resilience, guaranteeing data confidentiality and integrity, and granting access to data to authorised people (see Sec. 5.3 for security related challenges).

Data accessibility and resilience can be ensured by replicating data on the user's devices. Deciding how many copies are needed and where they should be placed depends on the kind of data and the context in which they are used. Some devices are only intermittently connected to the system, thus the data management service has to automatically adapt to the user behaviour and mobility to ensure on demand access of their data. Data copies may be created or moved, thus a location-independent naming scheme and content-based routing should be implemented to allow users retrieving data without knowing the exact location. For each kind of mutable personal data, a suitable data consistency model should be selected and implemented. Not all applications requires the most recent version of a data. Weak consistency models [26] should be used whenever appropriate because they require less resources and are more efficient than protocols for strong consistency.

While facing a tsunami of personal data, *garbage collection* is an important data management service as some data have a short life-time while others need to be kept longer. User defined policies regulates data time-to-live, thus this service should automatically delete all copies at time expiration [27].

Data management operations incur energy consumption, hence, the data management system needs to be designed in an energy-efficient way to cope with battery-powered devices.

5.2 Policies to enforce privacy

Privacy is the driving factor that motivates the design of our system. The first solution is to keep the data local in the personal cloud under the direct control of the user. Users may want to share with third parties some of their personal data, or data of other subjects under their responsibility, thus, a data access and usage control policy enforcement service has to be implemented [1]. This service will detect any access to data (create, delete, read, append, modify, transfer operation) and check if the requested access can be granted. There should be a simple way for users to express what services and users can access their data, for which purpose, in which context, and for how long. It should also be easy to revoke access or remove data wherever it is located. Policies are not static; users may want to add new policies and modify or remove existing ones.

A domain-specific language should be defined enabling the expression of policies that can be fine-grained, related to specific personal data items and individuals, or coarse grained, related to a large set of personal data or a group of individuals. Context is another important factor to take decisions on which data is relevant to share or on its sensitivity. Socially aware mechanisms could be designed enabling people to easily express their data management policies and helping them determine sensitivity indicators before sharing data [7].

Users should be able to tune their privacy and data management policies over time. Awareness of the amount of data manipulated and the accesses granted is highly desirable to improve users behaviour regarding privacy and data management, e.g. by implementing an audit service and notifications [28].

5.3 Security to enforce privacy

Identity management and efficient authorisation mechanisms are required to reduce the burden of the users who might need to access different services. Users have multiple accounts with different providers, making the management of their identities complex to handle. Federated identity could be a handy solution to eliminate the issue of managing multiple authentications. OpenID and OAuth [29] are two existing open standards, respectively allowing the use of third party services for authentication and providing secure delegation mechanisms to authorise third parties to access resources. *Level of assurance* could also be set to grant security levels and privileges on the resources, based on the trust in an Identity Provider (IdP) security mechanisms. It is worth mentioning the Moonshot project (www.project-moonshot.org) [30], a recent standard working group with the aim of providing federated identity at all levels, i.e., taking identities at the network level and bringing them to higher level services while preserving the privacy by not revealing personal information.

For guaranteeing data confidentiality and integrity, traditional encryption techniques can be used. However, cloud computing exacerbates the drawbacks of working with encrypted data, for instance in using stored data (queries and matching) or in processing data in the cloud. Several techniques have been proposed to obscure data in such a way that is still useful while maintaining privacy; for instance k-anonymity [31] makes the retrieved information indistinguishable within a given set. On the other hand, private information retrieval with oblivious transfer [32] can enforce privacy requirement at the user, who does not reveal what information she is accessing, and at the server provider, who does not disclose any additional one either. Differential privacy encryption [33] can be used to add a certain noise to data to protect private information while making the data still valuable in its aggregated form. Fully homomorphic encryption [34] has received considerable attention in the cloud computing community for outsourced processing (with no shared information on the input & output) since it allows to process encrypted data and get encrypted output. However, fully homomorphic encryption is not yet adopted due to its

high computational cost, high latency, and size overhead that makes it still inefficient for use in practice.

Despite the existing security primitives and tools to provide anonymity and confidentiality, viable security solutions should be investigated to guarantee privacy since the beginning. No assumption should be made on the trustworthiness of public cloud services; they should only be trusted under verification for storing and processing data, under user direct control.

5.4 Legislation to enforce privacy

Service Level Agreements (SLAs) are the primary contract form to regulate the interactions between end-users and service providers via the definition of service level objectives, including Quality of Protection (QoP) terms for privacy and protection constraints. While metrics can be used to monitor functional terms, QoP is hardly measurable; users have no means to verify if the provider has implemented specific security policies, e.g. secure deletion of data. Solutions are needed to allow users to verify, and not only trust, the compliance of a service provider with the QoP terms. In 2012, the Cloud Security Alliance has established the Privacy Level Agreement (PLA) working group (https://cloudsecurityalliance.org/research/pla) to define a set of best practises to check the compliance of service providers with data protection legislations. Europe has recently emanated the General Data Protection Regulation for the protection of personal data, and has promoted the use of Privacy Enhancing Technologies (PETs) to reduce the risk of privacy breach. However, the complexity in implementing security versus usability could mine the real adoption of PETs. Solving some of the challenges identified in previous paragraphs can help in this direction.

6 Concluding Remarks

In this paper we have focused on the overall design of a distributed system that follows a privacy by design approach exploiting locality to support users when using legacy cloud services, interacting, sharing information and collaborating with peers. We have identified and presented the technical challenges that should be addressed to implement the system, The attestation of the underlying platform running the virtualisation technology or the software itself [35] have not been discussed in this paper. While very important, they are not specific to the design of our system.

Another important aspect to consider is the incentives for the industry to adopt such an approach, despite not being a technical challenge. The benefits for the industry are both technical and commercial. Indeed, the Personal Cloud solution will reduce the load on the service provider infrastructure by leveraging the computation capabilities at the edge of the network, making possible to accommodate more users. At the same time, industry can offer a new service by renting instances to users for deploying their personal cloud on the public infrastructure. The provider could even advertise and sell, with additional costs, dedicated hosting services with privacy compliant certification, thus embracing a new set of users who were sceptical in using the service to avoid potential privacy risks. In all cases, the Personal Cloud solution could not be considered a substitute of public cloud providers for evident limitations in computation and storage capabilities, but as an important tool to create a user-centric ecosystem over undependable cloud providers. The user will still share data with external entities, but under direct control.

Future work includes designing an application programming framework well-suited to a new generation of applications taking advantage of the swarm of personal devices to offer people novel high-value services improving their day to day life, well-being, and enabling collective problem solving. We plan to study the impact of our system design on the integration of new devices,

and examine the deployment and upgrade of distributed applications and of the system services taking the standpoint of the device, application, and system services providers. To address scalability and help the system to keep function in case of disconnected access to the Personal Cloud, we envisage to investigate a distributed version. A combination of personal devices can be used to run the Personal Cloud management and security services.

Ultimately, we plan to further investigate the technical challenges presented in this paper and to implement a prototype to be experimented in a realistic environment with real users. The system we design is targeted to anyone owning personal devices and willing to use the emerging quantified-self, mobile social and crowd-* applications. We plan to collaborate with social scientists to better understand user perception of trust, privacy, system usability, and user motivation to use the afore-mentioned applications and system.

References

[1] S. Gnesi, I. Matteucci, C. Moiso, P. Mori, M. Petrocchi, and M. Vescovi, "My data, your data, our data: Managing privacy preferences in multiple subjects personal data," in *Proc. of Privacy Technologies and Policy*, ser. LNCS. Springer, 2014, vol. 8450, pp. 154–171.

[2] "NSA Prism program taps in to user data of Apple, Google and others," accessed on October 9th, 2014. [Online]. Available: http://www.theguardian.com/world/2013/jun/06/us-tech-giants-nsa-data

[3] "A european study on the nature of consumer trust and personal data," Orange, Tech. Rep., February 2014.

[4] "7 foundational principles – privacy by design." [Online]. Available: http://www.privacybydesign.ca/index.php/about-pbd/7-foundational-principles/

[5] "Rethinking personal data: Trust and context in user-centred data ecosystems," World Economic Forum, Tech. Rep., May 2014.

[6] "General Data Protection Regulation," European Commission, 2012.

[7] H. Harkous, R. Rahman, and K. Aberer, "C3P: Context-Aware Crowdsourced Cloud Privacy," in *Proc. of Privacy Enhancing Technologies*, ser. LNCS. Springer, 2014, vol. 8555, pp. 102–122.

[8] E. Rader, "Awareness of behavioral tracking and information privacy concern in facebook and google," in *Proc. of Symposium on Usable Privacy and Security (SOUPS)*, Menlo Park, CA, USA, July 9-11 2014.

[9] A. Khan, M. Othman, S. Madani, and S. Khan, "A survey of mobile cloud computing application models," *IEEE Communications Surveys Tutorials*, vol. 16, no. 1, pp. 393–413, First 2014.

[10] O. Hasan, B. Habegger, L. Brunie, N. Bennani, and E. Damiani, "A Discussion of Privacy Challenges in User Profiling with Big Data Techniques: The EEXCESS Use Case," in *Proc. of IEEE International Congress on Big Data*, Washington, DC, USA, 2013.

[11] J. Pan, S. Paul, and R. Jain, "A survey of the research on future internet architectures," *IEEE Communications Magazine*, vol. 49, no. 7, pp. 26–36, July 2011.

[12] A. Anand, F. Dogar, D. Han, B. Li, H. Lim, M. Machado, W. Wu, A. Akella, D. G. Andersen, J. W. Byers, S. Seshan, and P. Steenkiste, "XIA: An Architecture for an Evolvable and Trustworthy Internet," in *Proc. of 10th ACM Workshop Hotnets*, Cambridge, MA, USA, 2011.

[13] T. Anderson, K. Birman, R. Broberg, M. Caesar, D. Comer, C. Cotton, M. J. Freedman, A. Haeberlen, Z. G. Ives, A. Krishnamurthy, W. Lehr, B. T. Loo, D. Mazières, A. Nicolosi, J. M. Smith, I. Stoica, R. van Renesse, M. Walfish, H. Weatherspoon, and C. S. Yoo, "A Brief Overview of the NEBULA Future Internet Architecture," *SIGCOMM Comput. Commun. Rev.*, vol. 44, no. 3, pp. 81–86, Jul. 2014.

[14] V. Valancius, N. Laoutaris, L. Massoulié, C. Diot, and P. Rodriguez, "Greening the internet with nano data centers," in *Proc. of 5th ACM CoNEXT*, Rome, Italy, 2009.

[15] P. Ranganathan, "From microprocessors to nanostores: Rethinking data-centric systems," *Computer*, vol. 44, no. 1, pp. 39–48, Jan 2011.

[16] F. Jalali, R. Ayre, A. Vishwanath, K. Hinton, T. Alpcan, and R. Tucker, "Energy Consumption Comparison of Nano and Centralized Data Centers," in *Proc. of Greenmetrics.* Austin, Texas, USA: ACM, 2014.

[17] A. Lebre, J. Pastor, M. Bertier, F. Desprez, J. Rouzaud-Cornabas, C. Tedeschi, A.-C. Orgerie, F. Quesnel, and G. Fedak, "Beyond The Clouds, How Should Next Generation Utility Computing Infrastructures Be Designed?" in *Cloud Computing: Challenges, Limitations and R&D Solutions*, Mahmood, Zaigham, Ed. Springer, Nov. 2014.

[18] L. M. Vaquero and L. Rodero-Merino, "Finding your way in the fog: Towards a comprehensive definition of fog computing," *SIGCOMM Comput. Commun. Rev.*, vol. 44, no. 5, pp. 27–32, Oct. 2014.

[19] Z. Ou, B. Pang, Y. Deng, J. Nurminen, A. YI-Jski, and P. Hui, "Energy- and cost-efficiency analysis of arm-based clusters," in *Proc. of 12th IEEE/ACM CCGrid*, Ottawa, Canada, May 2012.

[20] F. P. Tso, D. R. White, S. Jouet, J. Singer, and D. P. Pezaros, "The Glasgow Raspberry Pi Cloud: A Scale Model for Cloud Computing Infrastructures," in *Proc. of IEEE 33rd ICDCS Workshops*, Los Alamitos, CA, USA, 2013.

[21] "Online labs – arm servers in the cloud," accessed on October 17th, 2014. [Online]. Available: http://labs.online.net

[22] D. Raychaudhuri, K. Nagaraja, and A. Venkataramani, "Mobilityfirst: A robust and trustworthy mobility-centric architecture for the future internet," *SIGMOBILE Mob. Comput. Commun. Rev.*, vol. 16, no. 3, pp. 2–13, Dec. 2012.

[23] K. Chard, S. Caton, O. Rana, and K. Bubendorfer, "Social cloud: Cloud computing in social networks," in *Proc. of IEEE 3rd CLOUD*, Miami, FL, USA, Jul. 2010, pp. 99–106.

[24] A. Mohaisen, H. Tran, A. Chandra, and Y. Kim, "Trustworthy distributed computing on social networks," *IEEE Transactions on Services Computing*, vol. 7, no. 3, pp. 333–345, July 2014.

[25] M. Coppola, Y. Jegou, B. Matthews, C. Morin, L. P. Prieto, O. D. Sanchez, E. Y. Yang, and H. Yu, "Virtual organization support within a grid-wide operating system," *IEEE Internet Computing*, vol. 12, no. 2, pp. 20–28, 2008.

[26] A. Bessani, M. Correia, B. Quaresma, F. André, and P. Sousa, "Depsky: Dependable and secure storage in a cloud-of-clouds," *ACM Transactions on Storage*, vol. 9, no. 4, pp. 12:1–12:33, Nov. 2013.

[27] J. Chow, B. Pfaff, T. Garfinkel, and M. Rosenblum, "Shredding your garbage: Reducing data lifetime through secure deallocation," in *Proc. of 14th USENIX Security Symposium*, Baltimore, MD, USA, 2005.

[28] R. Balebako, J. Jung, W. Lu, L. F. Cranor, and C. Nguyen, ""little brothers watching you": Raising awareness of data leaks on smartphones," in *Proc. of 9th ACM SOUPS*, Newcastle, United Kingdom, 2013.

[29] D. Hardt, "The OAuth 2.0 Authorization Framework," IETF RFC 6749, Oct. 2012.

[30] J. Howlett, S. Hartman, H. Tschofenig, E. Lear, and J. Schaad, "Application Bridging for Federated Access Beyond Web (ABFAB) Architecture," IETF Internet-Draft draft-ietf-abfab-arch-13.txt, Jul. 2014.

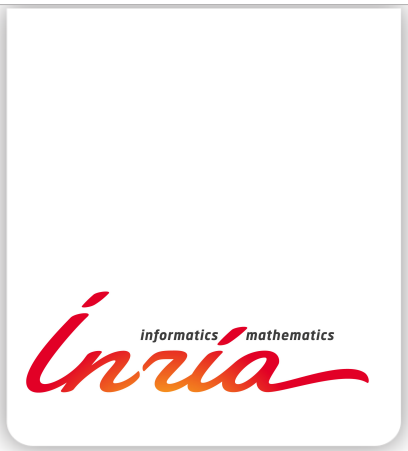
[31] V. Ciriani, S. De Capitani di Vimercati, S. Foresti, and P. Samarati, "k-anonymity," in *Secure Data Management in Decentralized Systems*, T. Yu and S. Jajodia, Eds. Springer US, 2007, vol. 33, pp. 323–353.

[32] J. Camenisch, M. Dubovitskaya, and G. Neven, "Oblivious transfer with access control," in *Proc. of 16th ACM CCS*, Chicago, IL, USA, 2009.

[33] C. Dwork, "Differential privacy," in *Proc. of 33rd ICALP conference*, ser. LNCS, vol. 4052, Venice, Italy, July 2006.

[34] C. Gentry, "Computing arbitrary functions of encrypted data," *Commun. ACM*, vol. 53, no. 3, pp. 97–105, Mar. 2010.

[35] M. Nanavati, P. Colp, B. Aiello, and A. Warfield, "Cloud security: A gathering storm," *Commun. ACM*, vol. 57, no. 5, pp. 70–79, May 2014.



RESEARCH CENTRE
RENNES – BRETAGNE ATLANTIQUE

Campus universitaire de Beaulieu
35042 Rennes Cedex

Publisher
Inria
Domaine de Voluceau - Rocquencourt
BP 105 - 78153 Le Chesnay Cedex
inria.fr

ISSN 0249-6399